

## WASHBACK OF ENGLISH NATIONAL EXAMINATION IN THE INDONESIAN CONTEXT

**Furaidah**

(furaidah02@gmail.com)

**Ali Saukah**

(alisaukah@yahoo.com)

**Utami Widiati**

(uwidiati@yahoo.com)

*Universitas Negeri Malang  
Jalan Semarang 5 Malang 65145, Indonesia*

**Abstract:** This study examines how teachers teach English to prepare students for high-stakes English national examination in the Indonesian context. Data were collected from two high-achieving and three low-achieving schools with eleven teachers as the subjects of in-depth interviews and non-participatory classroom observations. The findings reveal that bi-directional washback was found in both groups of schools. The schools of low achievers were found to have more intensive negative washback than those of high achievers. The different intensity of negative washback is likely related to the teachers' perspective about their students' level of competence for passing the national examination and about the quality of their schools. The consistently unsatisfactory national examination scores of the low-achieving schools, despite their concerted efforts in the examination preparation program, suggest that the government should focus on supporting such schools with more empirically-based empowerment programs, which would become an indispensable follow-up actions regarding the implementation of the high-stakes national examination.

**Keywords:** English national examination, washback, high-achieving schools, low-achieving schools, Indonesia

**DOI:** <http://dx.doi.org/10.15639/teflinjournal.v26i1/36-58>

In the Indonesian context, English is one of the subjects in junior and senior high schools, the final achievement of which is partly measured in the national final examination, as officially stated in the Ministry of Education Decree No. 34/2007. The result of the English National Examination (henceforth ENE) is used as one of the bases to determine student graduations from the schools and to consider for selection purposes at the higher levels of education. Such a decisive power of ENE classifies it as a high-stakes test that creates never-ending controversies.

In the local as well as national mass media, people have expressed agreements and disagreements towards that decisive role. The continuing pros and cons, however, have more frequently been based on personal opinions than on the findings resulting from empirical research studies. There have been very few studies on the effect of ENE as a high-stakes test on the teaching and learning process of English in Indonesian classrooms. The arguments also tend to be general as they are rarely directed specifically towards the effects of the national examination of a specific subject like English on the teaching and learning of the subject. Therefore, very little evidence can be used to support suggestions to improve the quality of the teaching and learning process of the subject.

Studies on the effects of a high-stakes test such as ENE can be found under the term 'washback' which, in this study, is defined as the effect of the national examination on the teaching and learning process. Several studies on washback reveal that the washback of a high-stakes test can actually be either negative or positive. Several studies on washback have been done abroad, such as in China (Qi, 2005), Hongkong (Andrew, Fullilove, & Wong, 2002), Jordan (Al Jamal & Ghadi, 2008), Taiwan (Chen, 2002), and USA (Stecher, Chun, & Barron, 2004).

The negative washback comes in many ways. The first is the refocus of teaching activities that result in the rearrangement of time allotment. More time is devoted to preparing students to take the test by spending more time for the teaching of the tested subjects (Chen, 2002; Ferman, 2004; Qi, 2005). Spratt's review of five studies on washback (2005) also shows that more curriculum time is spent on exam classes and that there are more students in exam classes than in regular ones. This reallocation of time is done at the expense of non-tested subjects, resulting in the sort of washback labelled as narrowing the curriculum or narrowing the scope and content of teaching and learning (Ferman, 2004; Qi, 2005). Spratt (2005) supports this finding as she also found the same

case from the five studies that she reviewed. Teachers teach the tested subjects, and once the test is over, they will pay attention to non-tested subjects again.

Changes in the content of teaching happen only superficially instead of being directed substantially to meet the achievement standards based on which a high-stakes test is constructed. It is superficial in that teachers only teach the parts that students will meet in a test. For example, as Stecher, Chun, and Barron (2004) point out, teachers focus on teaching students to write short paragraphs for the preparation to take the WASL (Washington Assessment of Student Learning) writing test. Such practices are in line with what Andrew, Fullilove, and Wong (2002) found. They reported that the sort of washback which is most apparent seems to represent a very superficial level of learning outcome, such as familiarization with the exam format and the rote-learning of exam-specific strategies and formulaic phrases. The use of contextually inappropriate phrases by a number of the students seems to indicate memorization rather than meaningful internalization of the tested language functions. Qi (2005) found that in preparing students for the writing section of a high-stakes university entrance test, the National Matriculation English Test in China (NMET), teachers focus more on linguistic accuracy, neglecting the communicative feature of writing a text.

Washback also appears in the teaching materials that teachers use. Spratt's review (2005) suggests that some teachers become textbook and exam slaves. In the former case, teachers rely heavily on textbooks, while in the latter, they rely even more heavily on past exam papers. Other teachers are reported to try innovative activities during exam preparation classes using a variety of self-made materials. Spratt concludes that teachers vary in using exam materials. An important factor related to this seems to be time; as the examination gets closer, the intensity of using past exam papers and commercial exam-related publications increases.

High-stakes tests tend to psychologically influence not only the teachers but also the students as they perceive the consequence of the test in their life. Washback also appears in the form of an increase in the teachers' and students' anxiety level (Ferman, 2004). Most of the teachers investigated admitted that the test aroused feelings of high anxiety and fear of test results. This was because the test results reflected their competence as teachers. Most students also reported that the test aroused in them feelings of anxiety to a quite high extent. The levels of ability tend to determine the levels of anxiety; weaker and aver-

age students are significantly more adversely affected by the potential failure in the test than the higher-level students are.

Besides the negative influences, high-stakes tests also make teachers and parents pay more attention to the students and children especially those with learning problems. Students also increase their intensity in learning. The increased attention seems to be positive; however, because the increased attention and involvement of the teachers, parents, and students do not always lead the teaching learning activities to the attainment of certain competences, such washback is also frequently deemed as negative as indicated by the following studies. Ferman (2004) reveals that teachers spend more time with their students and use the time to coach and drill them more intensively, which sometimes is done on individual basis. Obtaining high test scores is the target of all the teachers' attention. Students' learning interest earns more concern, and interestingly, students with higher level of ability are taught more deeply and broadly, aiming at the students' bigger possibility to reach higher scores from the test (Chen, 2002). It seems that the washback effect in the form of teachers' attention toward students is influenced by their perception towards the students' level of abilities. Ferman (2004) also reports that washback can also be seen in the parents' increased involvement toward their children's learning. Parents urged their children to study more seriously and even sometimes employed tutors to help their children prepare for the test.

Another positive effect of high-stakes tests is when the tests can direct teachers to teach the test to prepare their students; for example, when the format and content of the test are changed from multiple-choice questions to performance-based, teachers tend to teach more contextually. Stecher, Chun, and Barron (2004) who investigated the effect of the implementation of a statewide standard test called EARL (the Essential Academic Learning Requirements) and a writing test called WASL (Washington Assessment of Student Learning) found that teachers report changing their emphasis on some of writing topics, covering more writing behaviors as stated in the standard. Most of the teachers interviewed admitted increasing their emphasis on writing for different audiences, purposes, styles, and format, whereas considerably fewer teachers increased their coverage of writing conventions and writing process. Most of them also stated that the WASL contributed to the changes they made more than the EARL. Another study by Chen (2002) reveals that when a high-stakes test requires students to perform speaking activities, teachers also tend to teach more communicatively, shifting their teaching focus, from grammar-based

teaching to more communication-oriented teaching, promoting the learning of the oral skills and upgrading the oral skills. Additionally, Ferman (2002) reports that teachers also teach in a more integrated way, integrating listening and speaking to prepare students for oral examinations and give more opportunity to them to practice speaking. Similarly, Vogler (2002) also shows that changing multiple-choice questions to some or all constructed response questions encourages teachers to increase notably the use of open-response questions, creative/critical thinking questions, problem-solving activities, use of rubrics or scoring guides, writing assignments, and inquiry/investigation.

The previously cited studies show that high-stakes tests tend to influence the teaching and learning process in the classrooms negatively. However, although the unintended washback seems to be the corollary of high stakes tests, intended or positive washback is also present following the implementation of such tests. The positive washback was reported to happen with the implementation of high-stakes tests with more communicative or performance-based format and content.

In the Indonesian context, ENE is a high-stakes test with a decisive contribution to the total final scores of national examination or *Ujian Nasional* (henceforth *UN*) that determines the students' graduation from secondary schools. ENE has been implemented so far with a multiple-choice format. It has also been implemented with the discrepancy between what is tested with what is stated in the Standards of Competence (*Standar Kompetensi/SK*) and Basic Competences (*Kompetensi Dasar/KD*) of the Standards of Content, based on which teachers teach English. The *SKs/KDs* of the English subject require a level of achievements of the minimum communicative competence that covers the receptive and productive skills. On the other hand, *Standar Kompetensi Lulusan or SKL UN* (the table of specification of ENE), stated in the Appendix of Minister of Education's decree number 75/2009, shows that at senior high schools, ENE focuses on the reading and listening skills whereas at junior high schools, ENE focuses on the reading and writing. This *SKL UN* is the basis for the development of ENE items. With ENE as a high-stakes test implemented amid such condition, it is interesting to see how teachers prepare their students for ENE and how the teaching of English is affected by the preparation. Hughes (2003) states that if a test is regarded important, if the stakes are high, preparation for it can come to dominate all teaching and learning activities. This study is thus conducted to investigate what goes on in the Indonesian classrooms while teachers prepare their students for ENE. As the dominant

format of the high-stake tests was multiple-choice, the result of the study could possibly be concordant with the findings of the previous studies. In such case, the result of this research could provide more evidences about negative washback within the context of Indonesia. However, the writers expected some unique findings considering that student admittance to high schools in Indonesia made the schools relatively homogeneous in terms of the students' achievement. This is discussed further in the methodology section.

Prior to the present study, some preliminary observations on how teachers of senior high schools prepare their students for ENE were carried out at a school of high achievers (SHIGH) and a school of low achievers (SLOW). The categorization is based on the average score of junior high school national examination, which is used as the basis for student admission to senior high schools. The results of the observations revealed the changes that teachers made in preparing their students for ENE in terms of the instructional objectives or teaching focus, the materials used, the amount of time allocated, and the main classroom activities. These four aspects serve as the bases for the foci of the investigation that is focused on the twelfth grade, particularly in the second semester when the intensity of the preparation generally increased notably. This study thus aims at investigating how teachers teach English to prepare students for ENE particularly in terms of the focus of the teaching, the materials used, the allocation of time, and the main classrooms activities.

## **METHOD**

This exploratory qualitative study involved multiple settings, three SLOWs – schools of low achievers (coded as Senior High Schools– SMA *Sekolah Menengah Atas*A, D, and E) and two SHIGHs– schools of high achievers (coded as Senior High Schools– SMA *Sekolah Menengah Atas*B and C) located in Malang area, East Java Province, Indonesia. The use of such multiple settings was expected to allow the collection of more comprehensive data from a range of settings and thus led to confirmation of the findings. The number of the settings was determined by data redundancy, whereas the selection of the settings was based on the result of the preliminary observations. A rationale that high-stakes tests influence students of high and low achievers differently appeared as another theoretical basis for involving those differently-achieving schools.

The subjects of this study were eleven twelfth grade English teachers from the five schools involved in this study; they were coded as AK, AH, AR, BR, BS, CM, CB, DN, DB, DR, ES. All of the teachers hold bachelor degree in TEFL, and they are all experienced teachers with teaching experience of more than five years. In terms of their language competence, eight teachers have better language proficiency than the other three teachers, as indicated by their speaking performance during the teaching-learning activities and interviews.

Each teacher was observed at least twice. In the high achieving schools, some teachers were observed more than twice due to their varied teaching activities. The data about the phenomena under study, that is, the instructional objectives or teaching focus, the materials used, the amount of time allocated, and the main classroom activities were collected using non-participatory observations and one-on-one interviews.

The descriptive data collected were classified, and conclusions were drawn from the bottom-up data. The four foci of the investigation served as the preliminary classifications that generate more refined categories. Based on the data of both the settings and the phenomena, causal-logical relationship was established to render the explanation about the discovered phenomena and to suggest relevant propositions.

## **FINDINGS AND DISCUSSION**

### **Findings**

The findings show that washback of ENE indeed occurred in the teaching of English in the five schools under study, but in different degrees of intensity depending on the contexts, particularly in the areas of teaching focus, teaching materials, time allocation, and classroom activities as explained in the following.

### ***Teaching Focus***

In preparing their students for ENE, the teachers focused their teaching on listening and reading aligning the teaching with the *SKL UN (Standar Kompetensi Lulusan* or the table of specification of ENE) included in the policy of the Ministry of National Education Number 75/1995 on national examinations. The intensity of this focus was classified into intensive and less intensive. The

focus is said to be intensive when the teachers used almost all of the English teaching time for listening and reading drills. Conversely, it is considered less intensive when the teachers taught English as usual in the regular English time or did not change the regular English time into drilling activities. These drilling activities for listening and reading were only a part of their whole teaching and were carried out in the time scheduled for the drills. Generally, both the SLOWs and the SHIGHs gave intensive focus in Semester 2, when ENE is held, and less intensive in Semester 1. However, their detailed practice was different. The intensive focus was more likely to be found at the SLOWs whose input students' average score of junior high school national examination was below seven. The alignment with the *SKL UN* increased intensively in Semester 2 when the teachers at the SLOWs devoted the regular English teaching time and the added extra time to English for drilling classes concentrating on doing the listening and reading multiple-choice questions. The teachers, however, did not specifically design their teaching to teach the particular listening and reading skills included in the *SKL UN*. Instead, they assumed that by using previous ENE materials, those skills in the *SKL UN* were already covered. In the classroom interactions, the teachers did not always explain or discuss how to arrive at the answer to a certain question.

Within the teaching of listening and reading, the teachers teaching the twelfth grades at the SLOWs frequently discussed vocabulary. They generally said that their students had limited vocabulary. In the classroom interactions, both the teachers and the students raised vocabulary questions, but the teachers answered the questions more frequently. Most of the questions from the students were about the meaning of certain words, and the teachers usually directly gave the Indonesian equivalents. After the meaning of a word was given, the class continued discussing the next questions. The following is an excerpt quoted verbatim from a transcript on classroom interactions between a teacher, coded as AH, and her student, coded as S that depicts such a situation. The Indonesian words are basically the teacher's translation of the English words she used. Therefore, in such a case, the English equivalents are not provided. In some other cases, however, when deemed necessary, the English equivalents are provided.

#### Excerpt 1

AH : *Kemudian*\_\_\_ next, *ada kata-kata sulit*? Can you find another difficult word? Unaccounted for. *Tak terhitung*. Counted *bisa di \_ter\_hitung*.



- Unaccounted *tak terhitung*. Dari paragraph pertama ada kata-kata sulit?  
[Any difficult words in paragraph one?]
- S : recognition  
AH : recognition  
S : recognition  
AH : know \_\_\_\_ recognize..... (The teacher was writing the word 'recognize' and the synonym 'know' on the white board).  
AH : recognize *mengeta* \_\_ *hui* \_\_\_\_ know  
AH : *kemudian* next... paragraph *satu masih ada*? [Any other difficult words in paragraph one?]  
S : crowd  
AH : crowd *keramaian*.

Unlike the other teachers at the SLOWs, ES from SMA E frequently asked about parts of speech among the vocabulary questions. He frequently reminded his students to memorize synonyms. He gave his students a list of English words, their synonyms, and the Indonesian equivalents. The words were taken from the passages in the previous ENE packages. The following fragment describes classroom interactions between a teacher ES and his students in which ES asked about synonyms, a teaching activity that he frequently did. (Note: Code Ss means more than one student tried to answer ES's question.)

### Excerpt 2

- ES : Do you know the meaning of encourage?  
Ss : .....(silent)  
ES : Tita do you want to say something? Don't be afraid. Say something, come on.  
S : *memberi semangat*  
ES : Okay. Do you know the synonym? To motivate.  
ES : seek. The verb seek. All of you \_\_ seek.  
Ss : seek.  
ES : Anybody knows the meaning of seek? Okay Bitami, you want to say something? Say something. Don't be afraid. Say something.  
S : *mencari*  
ES : *men* \_\_ *cari*. What is the synonym of seek?  
Ss : look for  
ES : look for. Good. You remember them on the final exam.  
Ss : *amiiin*.

The increased focus was given only to reading and listening drills, and this indicates the teachers' tendency to teach to the test and narrow the curriculum or the *SKs/KDs* (the standard of competences/basic competences), which actually requires teachers to teach the four language skills. The teachers reported that even though the lesson plans submitted to the principals already covered the teaching of the four skills, the principal told them to focus on giving the listening and reading drills to anticipate ENE. The following excerpt taken from an interview between the researcher and AK, a teacher from SMA A, delineates such a situation.

**Excerpt 3**

- R : What did the school principal tell the subject teachers in preparing the students for ENE?
- AK : Usually the headmaster asks us to focus on the *SKL UN*, during ... what's that ... During especially the last two months before ENE. Focusing on that only. Not out of that.

At the SHIGHS, whose input students' average score of junior high school national examination was higher, the teachers also increased the teaching focus on listening and reading, but the alignment was not as intensive as that at the SLOWs. At SMA B, for example, the school provided drilling program in which the teachers drilled the students to do listening and reading questions. However, at the regular English teaching time, the teachers still taught all the four skills of listening, reading, speaking, though less writing. Data from observation and interview also suggest that at the regular English teaching in Semester 2, the students still read non-ENE-like materials, such as reading texts taken from newspapers or magazines and discussed the content in groups so that there were a lot of interactions among them. The students also had the opportunity to practice speaking by having presentation in front of their classmates and working in groups regardless of the effectiveness of the group work. One of the teachers, BS, explained that he frequently implemented contextual teaching and learning (CTL) in his classes.

Similar to SMA B, at SMA C the teachers taught reading and listening through drilling activities only in the scheduled drilling classes. Outside the drilling classes, the teachers taught as usual following the handouts developed by one of the teachers, CB. In CB's class, the students reviewed grammar (e.g., *causative have*) and practiced using it in new contexts, watched films, listened

to songs, and discussed the meaning of the song lyrics relating to their life. In these activities, the students used English in expressing their thoughts and opinions. CB said he taught the four skills in an integrated way. A similar communicative activity was also found in CM's class. Here, the students watched a film and wrote the review using their own words in a worksheet provided by the teacher. The worksheet helped them write a complete review because it contained elements of a review text that the students had to write about, such as the identity of the film, the orientation, and interpretative recount.

### ***Teaching Materials***

The findings related to teaching materials cover the kinds, reasons for selection, compression, and development. Each aspect is subsequently discussed in the following sub-headings.

#### **Kinds of Materials**

The teachers under study used ENE-like materials and non ENE-like materials. The ENE-like materials were those in the form of multiple-choice questions, such as previous ENEs, questions developed by the teachers or downloaded from the internet, and simulation tests from commercial publishers. The non ENE-like materials were in the forms of *BSE/Buku Sekolah Elektronik* (the government-endorsed electronic textbooks), imported textbooks, teacher-made handouts, or authentic reading and listening materials taken from newspapers, magazines, or downloaded from the internet.

The kinds of teaching materials used to prepare students for ENE appeared to concur with the teaching focus. At the SLOWs, in Semester 1, the teachers used *BSE* and ENE-like materials. In Semester 2, especially after the first month, with the shift of focus on listening and reading drills only, almost all of the materials that the teachers used were ENE-like materials. Non ENE-like materials were sometimes used just as variation. For example, at SMA A, only AK, the most senior English teacher at the school, who sometimes varied her teaching materials with reading texts with essay type questions. At SMA E, ES sometimes referred to *BSE* again to remind his students about certain information contained in the book, such as the characteristics of a certain text. However, most of the time, they used previous ENEs or other ENE-like materials.

With input students having higher scores of junior high school national examination, at the regular teaching time, even on the days closer to the D-day of ENE, the teachers at the SHIGHs still used various materials, ENE-like and non-ENE-like materials. At SHIGH, the multiple-choice questions were used only in the drilling classes. Outside the drilling classes, the teachers used non-ENE like materials, that is, reading and listening texts taken from many sources, such as the internet, imported books, magazines, newspapers, and teacher-made handouts. These handouts were in the form of electronic files covering the four language skills with a different language skill as a starting point for each chapter. For the drilling classes, the teachers used the worksheets consisting of multiple-choice questions.

### **Reasons for Material Selection**

Concerning the ENE-like materials, there were reasons for the selection which were common among all the teachers of both groups of schools, and which were common among the teachers of the SLOWs only. The collected data did not reveal any reason common just among the teachers of the SHIGH.

The reason shared by all the teachers was the suitability of the materials with *SKL UN*. In respect to the micro skills or the tested competences in *SKL UN*, such as getting main ideas or detailed information, some teachers explicitly said that they used assumptions to decide whether the materials were congruous with *SKL UN*. This explains why previous ENE papers tended to be the most frequently used teaching materials as the teachers believed in the congruence of previous ENEs with the *SKL UN*. Other teachers said that they examined the suitability with *SKL UN* by reading the questions before deciding to use previous ENEs. They admitted that if the questions they discussed did not correspond to the *SKLUN*, such materials were still worth using because the students could still take the benefit from reading them.

A reason shared by the teachers at the SLOWs only for the material selection was that the level of the difficulty of the teaching materials should not be beyond their students' competence. The source of difficulty frequently mentioned was vocabulary and this confirms their perception about their students' limited vocabulary repertoire. The following excerpt taken from an interview between the researcher (R) and AK shows how level of difficulty mattered in the selection of the teaching materials.

**Excerpt 4**

R : On what basis do you choose this material and not that material?

AK : In terms of the level of difficulty, suitable with *SKL UN*. Some vocabulary I have to adapt. I change into more simple and easy to understand for my students. At least the concept they have known. For me the concept is more important. For the vocabulary what's that \_\_\_\_\_ difficulties sometimes I give them and let them use dictionary.

The teachers at SLOWs added that the possibility of the questions and texts from previous ENE questions and from publishers' commercial simulation tests to appear in the upcoming ENE was another important criterion for material selection. Two teacher respondents acknowledged that some texts from previous ENEs and the simulation tests published by a certain publisher were indeed used several times in ENEs.

**Compression of Teaching Materials**

Most of the teachers at the schools involved in this study compressed their one-year teaching materials with different degrees of intensity and finished them sooner than scheduled. At the SLOWs, the compression was more intensive than at the SHIGHs. The common reason for the compression at both groups of schools was the exam-related events like try-out tests from *Diknas* (the local office of the Ministry of National Education and culture), school examination, and the shortened duration of the semester for the national examination. At the SLOWs, the intensity of the compression was the consequence of the intensive focus on listening and reading drills in Semester 2. At these schools, the teachers finished the teaching materials earlier, at the end of Semester 1 or at the beginning of Semester 2, and the following class activities were devoted to the listening and reading drills with ENE-like materials.

There were several ways of finishing the teaching materials ahead of the schedule. Skipping familiar topics, such as narratives, and reducing the discussion time and thus focusing more on reading and listening were the common practices done to compress the one-year teaching materials, especially at the SLOWs. At the SHIGHs the teaching materials were usually finished in Semester 2, but they were done more quickly due to many school examinations in that semester. The compression of the teaching materials was not intensive

because regular classes with non-ENE like materials still continued in Semester 2.

### ***Time Allocation***

In preparing the students for ENE, the schools increased the time allocation for all nationally-tested subjects including English. The addition of time allocation for English started in Semester 1 and increased in Semester 2. The amount of time added varied from school to school but, in general, the SLOWs allocated more time than the SHIGHs. The addition of the time ranged from three to five hours per week. Generally the added time was devoted to drilling and reviewing programs focusing on training the students in dealing with typical multiple-choice questions appearing on ENE.

### ***Classroom Activities***

The classroom activities were classified into regular teaching and drilling activities. It was called regular teaching activities when in the usual scheduled time for English the teachers and the students did not do or discuss ENE-like materials. In the regular teaching activities, the teachers held teaching learning activities that gave students more opportunities to use their language and interact with their classmates. In contrast, in the drilling activities, the students were only doing the ENE-like multiple-choice questions, including previous ENE questions. In other words, drilling activities were characterized with teacher-centered interactions. The teachers talked more often than the students did, and thus the interactions among the students were minimal.

The SHIGHs held the two types of activities while the SLOWs tended to focus more on one activity, that is, the drilling activities. The SHIGHs in general believed that their students already had good English competence. One of the teachers at the SHIGHs even explained that having good command of English was one of the requirements to be accepted at this school. Therefore, the English teachers did not seem to worry about not devoting their time to many drilling activities. Another teacher said that in her regular teaching she used more difficult passages and problems than those used in ENE questions, so she was sure her students would pass ENE. Students' boredom appeared to be another reason for not giving many drilling activities. At the SHIGHs it was very common that many students attended commercial drilling activities out-

side the school provided by commercial courses. When a question about the students' involvement in commercial drilling activities was given to a class, almost all students raised their hands indicating their participation. The teachers' statement about their students' boredom was also confirmed by few students coming to the drilling activities at the school.

At the SLOWs, where drilling activities appeared dominant, the classroom activities were geared toward only doing multiple-choice questions with some variety in the process from one school to another. There were two stages in doing the multiple-choice questions. The first was doing the questions, and the second was checking the answers together with the class led by the teacher. In the first stage, there were three patterns of activities. In the first pattern, the students read a reading text aloud, read the question (the stem), and then mentioned his/her answer from the options. In the second pattern, the teachers read a text and or the stem of a question aloud, called students in turn to choose the correct answer, and then the pointed student mentioned his/her choice. In the third pattern, the students were given a certain amount of time to read and do the questions individually.

The second stage of the drilling activities was checking time. The focus of this second stage was to check whether the students' choice answers were correct. The students usually mentioned their choices by reading the options (in the forms of sentences, phrases, or single words), or they just mentioned the letters (a, b, c, d) representing their choices. The teachers confirmed correct answers by showing their agreement (such as saying 'next' that means asking another student to continue with the next questions). Alternatively, the teachers directly asked the students to give the reasons for their opinion. The students gave the reasons by quoting relevant words/phrases/sentences from the texts as supporting evidence. When wrong answers were given, the teachers usually asked more students. When the pointed students or the class could not give correct answers, the teachers showed the clues in the texts, asked the students again, and explained the meaning of the sentences. If the students still could not give the correct answer, the teachers showed the correct choice.

In all of those activities, the students answered the questions individually and the classroom interactions were mostly between the teacher and the students. It was a question and answer kind of interaction with the teacher asking a question and students responding to the question by mentioning their choice answers. Most of the students' responses depended on the forms of the available options in the multiple-choice questions because in responding, the students

usually only read the provided options. The variation of their responses, thus, was limited to reading the sentences, phrases, words, or just mentioning the letters (a, b, c, d, e) that represented what was stated in the options.

## **Discussion**

The findings reveal that the teaching of English preparing the students for ENE indicates the occurrence of the washback of ENE. The different modes of carrying out the preparation suggest different intensity of washback of ENE at the SLOWs and at the SHIGHs.

At the SLOWs, the teaching of English, in terms of the focus, materials, time allocation, and activities, was almost totally aligned with the *SKL UN* of ENE. The stronger alignment with ENE at the SLOWs suggests that washback is more intensive at the SLOWs than at the SHIGHs, and it seems that only washback manifested in the increased amount of time allocation for UN subjects is likely to be beneficial as it can potentially support the development of communicative competence as stated in the standards of competences of the English subject.

The evaluation of whether washback is beneficial or harmful is usually based on the perspective of communicative language teaching. As Bailey (1996) argues, much of the concern about the alleged negative washback comes from the professionals' beliefs stating that standardized tests tend to be contrary to the principles and practices of communicative language teaching. In communicative language teaching, people negotiate for meaning in their attempts to understand and to be understood.

At the schools involved in this study, the positive washback was discerned in the more intensive attention given to English class as indicated by the provision of the examination preparation activities, or more particularly the significant increase in time allocation for English compared to the time allocation required by the Ministry of National Education and Culture regulation. At all of the observed schools, the allocation of time for English was higher than that required by the regulation, with the SLOWs allocating more time than the SHIGHs. The increase in time allocation in response to ENE is in line with Qi's (2005), Ferman's (2004), and Chen's (2002) findings. They find that in preparing students for a high stakes test, more time is devoted to the teaching of the tested subjects. In terms of attention, the increase in time allocation indicates the increase of the teachers' attention to students' learning, and this is



positive if none is sacrificed. For the teaching of English, this additional time should also be good news for it provides more time for learning and acquisition, and the development of the language competence required by the subject competence standards.

At the SLOWs, however, the rearrangement of time allocation was done at the expense of the other skills not tested by ENE. Washback on the teaching content was clearly indicated by the fact that the allotted time was merely spent on practicing answering listening and reading multiple-choice questions, disregarding the teaching of the other skills, speaking and writing. The phenomenon of teaching to the test that narrows the teaching content as found in this study confirms the criticism from the opponents of high-stakes testing policy as reported by Natriello (2009). The opponents assert that high-stakes testing policy triggers the occurrence of teaching to the test phenomena in which teachers only teach the parts to be tested. This present study is also in line with other studies revealing that such rearrangement of time allocation is done at the expense of other non-tested parts or subjects resulting in negative washback on teaching content. As previously mentioned, conclusion from studies by Ferman (2004), Qi (2005), Chen (2002), and Stecher, Chun, and Barron (2004) reveal that more time is devoted to preparing students for high-stakes test by spending more time for the teaching of the tested subjects. A research review by Spratt (2005) who examined five studies on washback also concludes that more curriculum time is spent on exam classes. The reallocation of time is done at the expense of other non-tested subjects, resulting in the sort of washback known as narrowing the curriculum or narrowing the scope and content of teaching and learning (Qi, 2005; Ferman, 2004). This situation affirms other research findings revealing that frequent changes in the content of teaching only happen superficially instead of being directed substantially to meet the achievement standards based on which a high stake test is constructed (Stecher, Chun, Barron, 2004; Andrew, Fullilove, Wong, 2002; Qi, 2005; Ferman, 2004).

As a consequence of teaching to the test (ENE) that narrows the teaching scope only to answering listening and reading question activities, washback is also found in the teaching materials, especially at the SLOWs. The results of this study are consistent with the findings of Spratt (2005), Cheng (1997), and Amengual-Pizzaro (2009). In her research review, Spratt (2005) concludes that some teachers become textbook slaves and exam slaves. As textbook and exam slaves, teachers rely heavily on textbooks and on past papers. As the exam gets closer, the intensity of using past papers and commercial exam-related publica-

tions increases. Cheng (1997), in her study on the washback of the Hongkong Certificate of Education Examination (HKCEE) in English in Hongkong secondary schools, reports that washback changes the teaching materials more quickly than the methodology that the teachers employ. With the launch of the more task-based test which required students' more integrated skills, nearly every school in Hongkong whose students would take the test used textbooks that had been revised to target the new examination. Cheng concludes that the teaching content receives the most intensive washback. Another study by Amengual –Pizzaro (2009) reveals a similar result. He reports that teachers use exam-related materials to get students fully prepared for the test, imitate the format of the test when making teaching materials, and use past papers widely. In conclusion, ENE exerted influence on teaching materials.

In addition, in both groups of schools, regardless of the amount of time, in the drilling classes, the teachers in general used similar ways of teaching. This is different from what Read and Hayes (2003) found in their study. Read and Hayes assert that time available affects teachers' choice of methodology. They found that with shorter preparation time period, teachers tended to use teacher-centered strategy while with longer time they allowed more participation from the students. In this study, at both the SHIGHS and the SLOWs, the classes were dominated by the practice of answering multiple-choice questions. The teachers talked more than the students did, and peer-interaction among the students was minimal because of the absence of group work and communication activities. Translating or giving Indonesian equivalents was opted for when the teachers dealt with words considered unfamiliar to the students. The similar teaching ways or methods at both groups of schools tend to be used due to the multiple-choice question format of the teaching materials.

In short, washback is discerned in both groups of schools, but in different degrees of intensity – being more intensive or stronger at the SLOWs than at the SHIGHS. This intensity is determined by the total alignment with ENE at the SLOWs and only partial alignment at the latter schools. The teachers' language competence, as indicated by their speaking performance in their classes and in the interviews, does not seem to be the factor contributing to the degree of the intensity of washback. The respondent teachers with good speaking competence are found at both the SLOWs and at the SHIGHS, and so are the teachers with lower speaking competence. All teachers at the SLOWs substituted their regular teaching almost totally with the drilling classes while the

teachers at the SHIGHs gave drilling classes only as additional classes to their regular teaching.

The length of the teachers' teaching experience does not seem to be another factor either. Senior teachers with longer teaching experience and junior teachers with shorter teaching experience are found at both groups of schools. Regardless of the length of their teaching experience, the teachers at the SLOW used drilling activities dominantly in preparing their students for ENE as opposed to the teachers at the SHIGH. Therefore, it can be concluded that teaching experience does not result in the different intensity of washback.

The factor that probably contributes more to the different intensity as found in the present study is likely to be the teachers' perception. Hughes (1993 in Bailey, 1996) explains that in the basic model of backwash (or washback) that consist of participants, process, and product, the teachers' perceptions (as participant) about the test can influence the teaching and learning task. Practicing the kinds of items predicted to appear in the test is an example of how the participants' perception influences the way they carry out their work or the process that will also affect the learning outcomes as the product of the work.

The teachers in this present study appear to be affected by ENE as a high-stakes test. That consequently affects any action taken by the participants that may contribute to the process of learning. The process includes materials development, teaching methodologies, etc. Bailey (1996) sees washback to teachers as washback to the program. The teachers' perceptions about their teaching responsibility to help the students pass the high-stakes ENE, however, are influenced by their perceptions about their students' level of English competence, which in turn determines the kind of teaching (process) taken to help the students. The kind of teaching is then customized according to the category of the participants. Bailey (1996) states that a test directly influences participants who are engaged in various processes and then results in product specific to each category of participants.

Another thing, not mentioned in the Spratt's review of research (2005), is that teachers do not act individually based on individual perspective or belief about their students' competence. Teachers as a community at a school have common perspective that influences what they do to prepare their students for the examination. At the schools involved in this study, the teachers were involved in the committee established to anticipate the national examination. The drilling activities were carried out not only by the twelfth grade teachers but al-

so by almost all teachers regardless of the grade they taught. Some of the respondent teachers were also assigned to collect or make the materials for the drilling classes.

To sum up, this study suggests that the washback of ENE seems to be more harmful at the SLOWs whose students are perceived by the teachers to have lower competence. To minimize the negative effect of ENE, the government should intervene the teaching of English at the SLOWs. The government should intentionally align ENE with the textbooks, inform this alignment to the teachers, and provide the guidance on how to carry out the intervention program. The implementation should also avoid the cramming fashion that teachers usually do when they prepare their students for ENE. As context-sensitive attitude is required from teachers, the teachers need to be equipped with knowledge and ability to decide the content based on the result of students' need assessment. In this way, though ready-to-use materials such as textbooks are at hand, the teachers will be able to use them judiciously and they will not be blindly dictated by the materials.

If for the time being, for some reasons, the intervention program to empower the schools with disadvantaged students cannot be implemented, it is suggested that the government reconsider the degree of the stakes ascribed to ENE. The reconsideration is important to guarantee more fairness to the students who are not ready for ENE because of the absence of an effective program at schools that can help them attain the level of competence sufficient to face ENE. The government should let the schools decide the degree of stakes that they want to take by considering their students' readiness.

To develop the students' readiness, the schools need to obtain data about what students need in terms of their competence. The government should encourage the schools to launch need assessment because so far the teachers rely solely on their perceptions about their students' competence. To confirm the teachers' perception about their students' current level of competence, a planned and systematic need assessment and follow-up action should be implemented. Such planned measure can promise more fruitful results than an instant solution taken only at the end of the students' academic life at senior high schools. A syllabus tailored to suit the students' distinct characteristics revealed by the need assessment needs to be developed to be the basis for the teaching of English preparing the students not only to succeed in the ENE but also, more substantially, to achieve the four language skills stated in the competence standards. The government can use the existence of need assessment and the

follow-up program as one of the criteria for the schools to get high grade in school accreditation to motivate the school to implement need assessment and the follow-up program.

## CONCLUSIONS AND SUGGESTIONS

In respect of the intensity of washback under study, that is, in terms of teaching focus, materials, time allocation, and classroom activities, this study reveals that teachers' perception about students' competence influences the degree of intensity. If teachers perceive that their students already possess the required competence to do ENE, washback appears less intensive. In other words, the more confident teachers are with their students' competence the less the degree of the washback.

However, the teachers' confidence or positive perception about their students can only form if their students indicate the level competence needed to pass ENE. This actually requires the teachers to look back at their three-year teaching that puts the twelfth grade students at the current competence *vis-a-vis* national examination. Whatever complaints the teachers have, such as the low level of the students' vocabulary size, which is the primary factor in verbal comprehension, the solutions cannot be done by cramming teaching in three or four months of examination preparation teaching. Cramming is ineffective because listening and reading skills (and other language skills) and vocabulary mastery develop incrementally.

Another noteworthy fact is that the intensity of washback is also related to the quality of the schools. Washback is less intensive at higher quality schools and more intensive at lower quality schools. As this study did not particularly investigate the quality of the schools involved in this study, this conclusion, however, needs further investigation particularly on how the daily teaching and learning process at both groups of schools is conducted to render a more valid conclusion.

At the schools with consistently unsatisfactory results, the government is suggested to do some intervention to help the teachers design and execute teaching programs that are theoretically more accountable in helping the students achieve the required competence. If the students can achieve the competence, passing ENE satisfactorily will not be a big problem. The students need all help to arrive at the ideal situation. In fact, one of the purposes of the im-

plementation of national examination is for the government to get data based on which the government can issue contextually appropriate policies.

Research also needs to be done at the schools whose result of ENE is consistently low. It is important to investigate how the daily teaching of English is held to find out the quality of the teaching process and the possible causes of unsuccessful teaching indicated by the consistently lower scores. This study only investigated the teaching of English in the second semester of the twelfth grade. A study with more complete coverage that includes the daily teaching and learning process in the tenth and eleventh grades will yield more complete data about the sources of the problems that hamper the students from achieving satisfactory ENE scores. This is particularly important at the schools that allocate much more hours than the national requirement. The problem to pursue can be how the allocated time is utilized to teach English and whether the teaching of English at the schools already meets the principles of the teaching of English as a foreign language.

## REFERENCES

- Amengual-Pizzaro, M. (2009). Does the English test in Spanish university entrance examination influence the teaching of English? *English Studies*, 90(5), 582-598.
- Al-Jamal, Du & Ghadi, N. (2008). English language general secondary certificate examination washback in Jordan. *Asian EFL Journal*, 10(3), 158-186.
- Andrew, S., Fullilove, J.& Wong, Y. (2002). Targeting washback – A case study. *System*, 30(2), 207-223
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257–279.
- Chen, L. M. (2002). *Washback of a public exam on English teaching*. Retrieved from ERIC Database. (ED 472167)
- Cheng, L. (1997). *The washback effect of public examination change on classroom teaching: An impact study of the 1996 Hongkong certificate of education in English on the classroom teaching of English in Hongkong secondary schools*. (Doctoral Dissertation, University of Hongkong).

- Ferman, I. (2004). The washback of an EFL national oral matriculation test to teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research context and methods* (pp. 191-210). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Natriello, G. (2009). High stakes testing and teaching to the test. In L. G. Saha, & A. G. Dworkins (Eds.), *International Handbook of Research on Teachers and Teaching*. Springer Science+Business Media, LLC.
- Qi, L. (2005). Stakeholder's conflicting aims undermine the washback function of a high-stake test. *Language Testing*, 22(2): 142-173.
- Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9(1), 5-29.
- Stecher, B., Chun, T., & Barron, S. (2004). The effects of assessment-driven reform on the teaching of writing in Washington State. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research context and methods* (pp. 191-210). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Vogler, K. E. (2002). The impact of high-stakes, state-mandated student performance assessment on teachers' instructional practices. *Education*, 123 (1), 39-51.